4
Iterative Reconstruction Methods

B.F. Hutton*, J. Nuyts§ and H. Zaidi†

1. Introduction

The direct reconstruction in emission tomography (ET) using analytical methods was described in chapter 3. This included filtered back projection (FBP), until recently the most commonly used method of reconstruction in clinical practice. Analytical reconstruction usually assumes a relatively simple model of the emission and detection processes and would become quite complex if rigorous models were applied. FBP has further limitations due to the presence of streak artefacts that are particularly prominent near hot structures and the noise enhancement that is inherent in the reconstruction. An alternative to analytical reconstruction is the use of iterative reconstruction techniques, which can more readily incorporate more complex models of the underlying physics and also can better accommodate assumptions regarding the statistical variability of acquired data. Unlike analytical reconstruction where FBP dominates, there are many approaches to iterative reconstruction. Difficulties in understanding the classification of these algorithms have led to considerable confusion in the choice of appropriate algorithms, particularly in routine clinical application. This chapter is intended to provide a general overview of iterative reconstruction techniques with emphasis on practical issues that may assist readers in making an informed choice. Iterative reconstruction can be applied equally well to single-photon emission computed tomography (SPECT) or positron emission tomography (PET) or indeed any tomographic data. The coverage in this chapter will be general, although some specific issues relating to either PET or SPECT will be highlighted. Reconstruction of transmission data will also be briefly addressed. For additional coverage readers are referred to other recent texts that specifically address iterative reconstruction.[1-3]

*Prof. B.F. Hutton, Institute of Nuclear Medicine, University College London, London WIE 6BT, UK; also Centre for Medical Radiation Physics, University of Wollongong, NSW 2522, Austalia
§Dr J. Nuyts, Katholieke Universiteit Leuven, Nuclear Medicine Department, U.Z. Gasthuisberg, Herestraat 49, B30000 Leuven, Belgium
†PD Dr H. Zaidi, Division of Nuclear Medicine, Geneva University Hospital, 1211

2. What is Iterative Reconstruction?

2.1 General Iterative Techniques

Iterative techniques are common in problems that involve optimization. The reconstruction problem can be considered a particular case where one is trying to determine the 'best' estimate of the distribution of activity based on the measured projections. An assumption underlying analytical reconstruction techniques is that there is a unique solution. However in practice, due to presence of noise, there are normally a number of possible solutions to the reconstruction problem. The 'best' reconstruction is determined by defining some criterion that measures goodness of fit between the reconstruction estimate and measurements and by adopting an algorithm that finds the optimal solution. Iterative techniques are well suited to solving this sort of problem.

A flow-chart that illustrates the general iterative technique used in reconstruction is presented in Figure 1. The iterative algorithm involves a feedback process that permits sequential adjustment of the estimated recon-

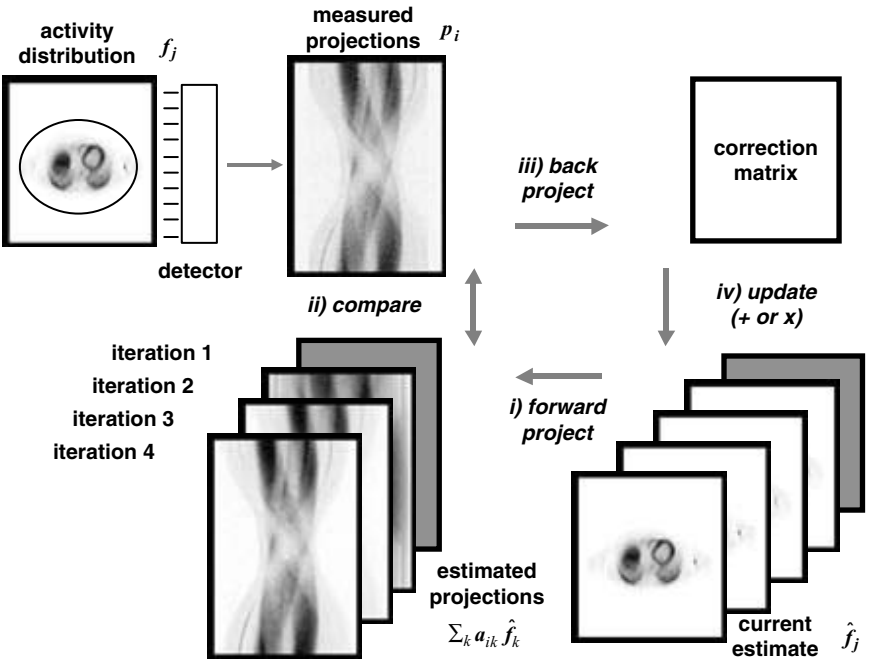

FIGURE 1. Schematic of general iterative reconstruction algorithm. Starting from a uniform grey image, estimated projections are constructed by forward projection. These are compared with the measured projections and a correction matrix is constructed by back projection. The reconstruction estimate is update either by summation or multiplication and this becomes the starting point for the next iteration.

struction so as to improve its correspondence with the measured projections. The iterative process begins with an initial estimate of the object count distribution, which may either be simply a uniform arbitrary count level or a previous reconstruction (e.g. using FBP). Provided a suitable model of the emission and detection physics is incorporated, the projections that would arise from this initial object can be estimated (by forward projection): this is effectively an estimate of what the detectors would measure given the initial object. Clearly this estimate of projections will differ from the actual measured projections, unless the initial guess happens to coincide exactly with the actual object. The difference between estimated and measured projections can be used to modify the original estimate of the object by use of suitable additive or multiplicative corrections at each point (usually via back projection). The adjusted object then becomes the starting point for a second iteration. This proceeds as the first with the forward projection so as to re-estimate the projections that would derive from this updated object. The process continues for multiple iterations using the continuous feedback loop until a final solution is reached (where usually a predetermined number of iterations have been completed, resulting in very little object change between iterations). This iterative process is central to all iterative reconstruction algorithms.

2.2 Attractions of Iterative Methods

Iterative reconstruction has a number of potential advantages that make it attractive in comparison with analytical methods. Foremost is the limiting assumption in analytical techniques that the measured data are perfectly consistent with the source object, a requirement that is never true in practice given the presence of noise and other physical factors (e.g. attenuation). The lack of consistency generally results in problems with noise and possible problems in finding a unique analytical solution. In contrast, iterative techniques function with either consistent or inconsistent data and can attempt to model noise directly, with improved noise properties in the final reconstruction. In addition iterative algorithms are well suited to handling complex physical models of the emission and detection processes, including position variant attenuation or distance dependent resolution. This ability to directly model the system, including some consideration of the noise characteristics, provides considerable flexibility in the type of data that can be reconstructed.

Iterative reconstruction methods have various additional properties that make them attractive, although some properties (e.g. non-negativity) are specific to the particular type of reconstruction and will be dealt with in the appropriate section. In general though, iterative methods provide reconstructions with improved noise characteristics. They greatly reduce the streaking artefacts that are common in analytical reconstruction and are better able to handle missing data (such as may occur due to truncation or inter-detector gaps). They generally provide accurate reconstructions that

can be used for quantification, provided some caution is exerted in choice of reconstruction parameters.

The main limitation in using iterative algorithms is the execution speed, which is significantly slower than FBP. However improvements in computer speed in combination with acceleration techniques (described in section 7) permits reconstruction in a clinically acceptable time.

2.3 Classification of Iterative Reconstruction Methods

It is convenient at the outset to introduce a general classification for the various iterative reconstruction methods. These broadly form three logical groups, defined mainly by the underlying assumptions regarding the nature of the data, whether statistical or non-statistical. The earliest developments involved no assumptions regarding the origins of noise in emission data and led to the various classical algebraic reconstruction techniques, which are non-statistical. These are described in section 3. The statistical techniques can be broadly divided into two main groups, those that assume Gaussian noise (involving least squares solutions: section 4) and those that assume Poisson noise (involving maximum likelihood solutions: section 5). These two fairly basic assumptions have led to two distinct classes of reconstruction that involve different algorithms and different properties. Variations on these algorithms specifically address issues such as noise control and acceleration (described in sections 6 and 7 respectively).

2.4 Nomenclature for Discrete Iterative Reconstruction

It is useful to establish nomenclature with reference to the general iterative algorithm in Figure 1. Consider the measurement of a set of projections $p(s,\phi)$, where s represents distance along a projection and ϕ represents the angle of acquisition, originating from a distribution of activity in the corresponding single slice in an object given by $f(x,y)$. It is convenient to consider the equivalent discrete set of projection pixel values p_i for counts originating from the object voxel activity concentration f_j. The process of estimating projections from a known (or estimated) activity distribution requires a description of the physical model that adequately describes the emission and detection processes. This essentially defines the probability of detecting an emitted photon, originating from location j, at any particular position, i, on the detector, which will depend on many factors including detector geometry, attenuation and resolution. A single matrix, a_{ij}, referred to as the transition matrix or system matrix, can conveniently describe this relationship. The process of forward projection can then simply be expressed as

$$p_i = \sum_j a_{ij} f_j \tag{1}$$

Similarly the back projection operation can be expressed as:

$$f_j = \sum_i a_{ij} p_i \tag{2}$$

3. Algebraic Reconstruction Algorithms

Algebraic reconstruction techniques are based on relaxation methods for solving systems of linear equalities or inequalities. Recent developments suggest that with proper tuning, the convergence of these algorithms can be very fast. However, the quality of the reconstructions has not been investigated thoroughly. In this section, a summary overview of algebraic reconstruction techniques is first set out, followed by a description of the algorithmic implementation of the classical algebraic reconstruction technique (ART) and its variants.

3.1 ART

The ART algorithm is based on the Kaczmarz method and uses a successive over-relaxation method well-known in numerical linear algebra.[4] It was first proposed as a reconstruction method in 1970 by Bender *et al.*[5] and was applied mainly to X-ray photography and electron microscopy. The early developments addressed mainly reconstruction of X-ray CT data but more recently, ART-type algorithms have been proposed specifically for PET.[6] The principle of the basic ART algorithm consists of describing every iteration point by point (row-action method), and in correcting all voxels in the image which are found on a projection ray, so as to minimise the difference between the values of the calculated and measured projections at the point under consideration. The process comes to a stop when a certain criterion becomes relatively small. For example the sum of squared differences between the calculated and measured projections can be used as such a criterion.

In fact, ART consists of guessing at a value for all the voxels f_j, and then modifying each element along each ray by a factor which compensates for the discrepancy between the measured and the calculated ray sum:

$$f_j^{new} = f_j^{old} \frac{p_i}{\sum_k a_{ik} f_k^{old}} \tag{3}$$

Here f_j^{new} and f_j^{old} refer to the current and previous estimates of the reconstructed object respectively. If the calculated ray sum is the same as the measured value, it implies that the guessed value is correct for a particular projection; however, for another projection there might be a large discrepancy. Thus the pixels of the last views (while lying in the ray for the new view) will be modified according to the discrepancy between the new ray and the measured value. Thus, each ray from each projection is examined and values of f falling within that ray are changed iteratively for all the

projections for a certain number of iterations. It is evident that the computational effort of the ART algorithm is relatively low, but data matrices can be several hundred of Megabytes in size.

It has been shown that by careful adjustment of the relaxation parameters and the order in which the collected data are accessed during the reconstruction procedure, ART can produce high-quality reconstructions with excellent computational efficiency.[4] In particular, the choice of the projection data access scheme proved to be crucial for improvement of low-contrast object detection.[7] An important modification of ART consists of setting to zero those values in the array that are clearly zero because they correspond to a ray sum that was observed as zero. This is an important boundary condition for any of the iterative techniques.

3.2 Other Variants

A large number of ART variants have subsequently been proposed in literature. Eq. (3) above is called multiplicative ART (MART). Another method of correcting the discrepancy between the measured and calculated projections consists of adding the difference between them. This is called the additive form of ART (AART). The diverse variants of ART correspond to row-action techniques, as each iteration only calls for the use of one equation at a time. Other algorithms like the simultaneous iterative reconstruction technique (SIRT) consist of correcting simultaneously each voxel for all rays passing through it. So, these corrections are incorporated by using data from all of the projections simultaneously. The simultaneous version of MART (SMART) is similar to the ML-EM approach (see section 5.1).

The block-iterative version of ART (BI-ART) is obtained by partitioning the projection set and applying the ART equation to each subset (a block represents a subset of the projection data). The correction is computed for each block, and may be interpreted as a weighted back-projection of the difference between the computed image projection and the acquired projection. This correction is performed only after an entire projection image is computed. Other variants of ART include the block-iterative version of SMART (BI-SMART) and the rescaled block-iterative SMART (RBI-SMART).[8] With BI-SMART in PET, an issue is the way to define a block to ensure the convergence of the algorithm, at least for the consistent case.

4. Statistical Algorithms-Gaussian Assumption

4.1 Least Squares Solutions

In nuclear medicine, the count rates are usually fairly low, and as a result, the data tend to be rather noisy. As mentioned above, one cannot hope to recover the true solution from such data, and some criterion to define the

'best solution' is needed. In statistical reconstruction, the best solution is defined as the one that is most likely, given the data. Thus, the reconstruction is obtained by finding the image $\hat{f}$ that maximises the conditional probability prob$[f|P]$, where P are the measured projections. Using Bayes' rule, this objective function can be rewritten in a more convenient form:

$$\text{prob}[f|P] = \frac{\text{prob}[P|f]\text{prob}[f]}{\text{prob}[P]} \tag{4}$$

Because P is constant, maximising prob$[f|P]$ is equivalent to maximising prob$[P|f]$prob$[f]$. The factor prob$[P|f]$ is called the "likelihood" and tells how well the data agree with the image. The factor prob$[f]$ is the "prior", and tells what is known already about the image f, prior to the measurement. Finally, prob$[f|P]$ is called the "posterior", and represents what is known by combining the a-priori knowledge with the information obtained from the measurement. If it is assumed that nothing is known in advance, then the prior is constant and maximum-a-posteriori (MAP) reconstruction is equivalent to maximum-likelihood (ML) reconstruction. The discussion of the posterior is deferred to a later section.

In emission and transmission tomography, the noise on the counts measured at two different detector locations is uncorrelated (the noise is "white"). This allows for factorisation of the likelihood:

$$\text{prob}[P|f] = \prod_i \text{prob}[p_i|f] \tag{5}$$

Maximising a function is equivalent to maximizing its logarithm. Therefore, the maximum-likelihood reconstruction $\hat{f}$ is obtained by maximising

$$\text{log-likelihood} = \sum_i \ln \text{prob}[p_i|f] \tag{6}$$

Assuming that the noise can be well approximated as a Gaussian distribution with known standard deviation makes the maximum-likelihood solution identical to the least squares solution. The mean of the Gaussian distribution for detector i is computed as the projection of the image $\hat{f}$. Consequently, the least squares solution is obtained by minimizing (over F):

$$L_G(P,F) = \sum_i \frac{\left(p_i - \sum_j a_{ij} f_j\right)^2}{2\sigma_i^2} = \frac{1}{2}(P-AF)'C^{-1}(P-AF) \tag{7}$$

where P is a column matrix with elements p_i, A is the system matrix with elements a_{ij} and F is a column matrix with elements f_j, and prime denotes transpose. C is the covariance matrix of the data, which is assumed to be diagonal here, with elements $c_{ii} = \sigma_i^2$. A straightforward solution is obtained by setting the first derivative with respect to f_j to zero for all j. This solution is most easily written in matrix form:

$$\hat{F} = \left[A' C^{-1} A\right]^{-1} A' C^{-1} P \tag{8}$$

The problem with this expression is the large size of the matrices. The inverse of the so-called Fisher information matrix $A' C^{-1} A$ is required.[9] It has $J \times J$ elements, where J is the number of voxels in the image to be reconstructed. Its matrix elements equal

$$\text{FIM(j,k)} = \sum_i \frac{a_{ij} a_{ik}}{\sigma_i^2} \tag{9}$$

which is mostly non-zero: although the diagonal dominates, this matrix is not sparse at all and direct matrix inversion is prohibitive.

As an exercise, it can be assumed (although totally unacceptable) that all standard deviations σ_i are identical and equal to 1. Moreover, consider an idealized parallel-hole projection, with perfect resolution and no attenuation or other degrading effects. The image minimizing L_G then becomes

$$\hat{F} = [A' A]^{-1} A' P \tag{10}$$

The operator $A' A$ computes the backprojection of the projection of a point. In this idealized case, the operator $A' A$ is shift invariant, and it can be shown that its inverse is (a digital version of) the inverse Fourier transform of the ramp filter. Consequently, this analysis reveals that FBP computes the solution of the unweighted least squares problem.

In emission tomography, it is not realistic to assume that all the standard deviations σ_i are identical. As a result, the Fisher information and its inverse become more complex, position dependent operators, that cannot be implemented with shift invariant filters (such as a ramp filter). Consequently, one has to turn to numerical, iterative optimisation techniques.

Another problem is the estimation of the standard deviations σ_i. The Poisson distribution is well approximated as a Gaussian with variance equal to the mean. But the mean (noise-free) projection count is unknown. There are two approaches to deal with this problem. The first one is to estimate σ_i from the noisy counts p_i. One could simply set $\sigma_i = {p_i}^{0.5}$. However, this leads to a negative bias in the reconstruction: all counts that happen to be smaller than the mean will be assigned a smaller standard deviation, and hence a higher weight in the weighted least squares computations. These noisy weights may even cause streak artefacts in the reconstruction. So the noise on the weights has to be reduced, e.g. by estimating σ_i from a smoothed version of the data. The second approach is to estimate σ_i from the current reconstruction, rather than from the data. This is closer to the physical reality, because the measurement is assumed to be a noise realization of the noise-free projection of the tracer distribution. However, it is also more complex, because now the weights depend on the unknown solution of the problem. Possibly because of this, the data-based estimate of the weights has received more attention in the literature.

4.2 Approaches to Minimization

Many optimisation algorithms can be applied to this problem (e.g. conjugate gradient or steepest descent), but it is beyond the scope of this text to present these in detail. One thing that most of these algorithms have in common is that they will modify each voxel such as to decrease the value of L_G in every iteration. To do that, the value added to voxel j must have the opposite sign to the derivative of L_G with respect to f_j. This can be written as

$$\hat{f}_j^{new} = \hat{f}_j^{old} - \frac{1}{\alpha_j} \frac{\partial L_G}{\partial f_j}\bigg|_{\hat{f}_j^{\text{old}}} \tag{11}$$

where α_j is some positive value, that may change during iteration, and that has to be tuned in order to guarantee convergence. To show the typical form of the optimisation algorithms, the derivatives of the two following cost functions are computed:

$$L_{G1}(P,F) = \sum_i \frac{(p_i - \hat{p}_i)^2}{2\sigma_i^2} \text{ and } L_{G2}(P,F) = \sum_i \frac{(p_i - \hat{p}_i)^2}{2\hat{p}_i} \text{ with } \hat{p}_i = \sum_j a_{ij}\hat{f}_j \tag{12}$$

L_{G1} is the cost function with pre-computed (data based) standard deviation. L_{G2} is the version where the variance is estimated from the current reconstruction. The derivatives with respect to f_j equal:

$$\frac{\partial L_{G1}}{\partial f_j} = -\sum_i a_{ij} \frac{p_i - \hat{p}_i}{\sigma_i} \text{and} \frac{\partial L_{G2}}{\partial f_j} = -\sum_i a_{ij} \frac{(p_i - \hat{p}_i)(p_i + \hat{p}_i)}{2\hat{p}_i^2} \tag{13}$$

The difference between algorithms is mainly in how the step size α_j is tuned. Note that (weighted) least squares algorithms do not "naturally" produce non-negative solutions. If non-negativity is desired, it must be imposed during the iterations, which can adversely affect convergence. Sauer and Bouman[10] and Fessler[11] proposed an effective method, which has been used successfully by others.[12,13] It updates the reconstruction voxels sequentially (as in ART), sets the step size α_j equal to the element *FIM(j,j)* of the Fisher information matrix and produces a non-negative solution. Other solutions have been proposed, e.g. scaled steepest descent by Kaufman[14] and a preconditioning algorithm by Chinn and Huang.[15]

5. Statistical Algorithms-Poisson Assumption

5.1 Maximum Likelihood Solution: the ML-EM Algorithm

The random nature of radioactive decay suggests that a Poisson model is more appropriate for emission data (although this is well approximated by a Gaussian provided measured counts are reasonably high). An appealing

consequence of using the Poisson model is that non-negativity is assured even at low count levels.

The basic Poisson model provides the probability of measuring a particular count, c, given an expected measurement, r:

$$\text{prob}[c|r] = \frac{e^{-r}r^c}{c!} \tag{14}$$

Using this Poisson model the probability of acquiring the projection count distribution that was measured, P, given an estimated distribution of activity in the emission object, f, can be represented by the product of probabilities for individual projection pixels. This conditional probability is referred to as likelihood, L

$$L(P|f) = \text{prob}[P|f] = \prod_i \exp\left[-\sum_j a_{ij} f_j\right] \left(\sum_j a_{ij} f_j\right)^{p_i} (p_i!)^{-1} \tag{15}$$

As argued above (section 4.1, eq (4)-(6)), maximising (the logarithm) of prob$[P|f]$ is equivalent to maximising prob$[f|P]$, provided the *a priori* distribution can be assumed to be constant. This provides the most likely distribution of emissions that represents the original activity distribution, given the measured projections.

There are various approaches to determine the maximum likelihood (ML) solution but the most commonly used is the expectation maximization (EM) algorithm, which unifies various previous statistical approaches.[16] This involves an iterative process with the attractive property that convergence is guaranteed. ML-EM was originally applied to emission tomography in the early eighties[17,18] but continues to be widely used. Other groups also deserve credit for much of the early application of ML-EM to emission tomography (e.g. Miller *et al.*[19]). The EM algorithm involves two distinct steps. First the expected projections are calculated by forward projection using the appropriate system / transition matrix, based on the estimate of the activity distribution from the previous iteration (an initial guess in the case of the first iteration). Second the current estimate is updated so as to maximise the likelihood, achieved by multiplication of the previous estimate by the back projection of the ratio of measured over estimated projections. The resultant ML-EM equation is derived elsewhere, as cited above, and is given by

$$f_j^{new} = \frac{f_j^{old}}{\sum_l a_{lj}} \sum_i a_{ij} \frac{p_i}{\sum_k a_{ik} f_k^{old}} \tag{16}$$

5.2 Properties of ML-EM Reconstruction

The EM algorithm results in an iterative process for estimation consistent with the general iterative flow-chart given in Figure 1. In this case the update is multiplicative, not unlike SMART, with the update simply being the ratio

of measured over estimated projections. As the number of iterations increases the likelihood increases, providing an estimate that theoretically is more likely to be close to the true object distribution. In practice, however, the image reaches an optimal visual quality at typically around 16 iterations and, in the absence of any noise constraint, appears progressively more noisy at higher number of iterations (Figure 2). The noise characteristics are appealing, with the variance remaining proportional to number of counts rather than being approximately position-independent as in FBP.[9,20] This tends to favour lesion detection in low count areas where the signal to noise ratio can be markedly improved. The ML solution after a large number of iterations is not the most 'desirable' solution as it reflects the actual noisy distribution of emitted counts rather than the underlying activity, whose distribution is likely to be much less variable. In clinical practice it is common to stop at a small number of iterations in order to limit noise. However it should be recognised that the reconstruction does not converge at the same rate for all points. Halting the reconstruction early runs a risk of reducing reconstruction accuracy, which can be avoided by using a larger number of iterations with post-reconstruction smoothing.[21] Alternative approaches to controlling noise are discussed in section 6.

There are several attractive theoretical properties of ML-EM although in practice these rarely offer real advantage due to approximations in the system model and the relatively high level of noise in most emission tomography studies. The use of a multiplicative update guarantees positive values and also means that areas outside the object, where zero counts are expected,

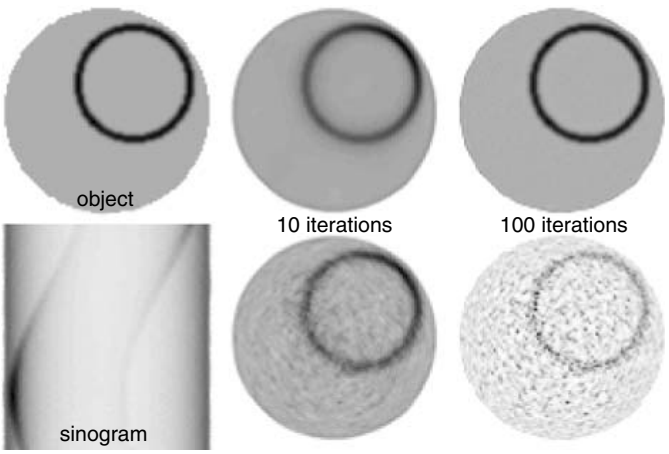

FIGURE 2. PET simulation of circular object in a uniform background, with uniform attenuation. ML-EM reconstructions at 10 and 100 iterations are shown, obtained from a sinogram without noise (top) and with Poisson noise (bottom). At 10 iterations, convergence is clearly incomplete. At 100 iterations, the noise level is becoming unacceptable.

are constrained so as to remain zero. This can result in problems near object edges when the projection counts are low as any zero pixel on a projection defines a projected line where counts are set to zero and remain unchanged, even when intersected by non-zero projections. The conservation of counts that occurs, results in the possibility of misplaced counts (usually hot spots) near the object edge.

A definite limitation with ML-EM reconstruction is the time taken for reconstruction. Each of the forward and back projection steps in all iterations takes approximately the same time as FBP; twenty iterations ML-EM would therefore take forty times as long as FBP. Fortunately steps can be taken to significantly accelerate the reconstruction as outlined in section 7.

6. Approaches to Controlling Noise

As mentioned above, the ML-criterion does not prevent noise propagation, and in many clinical cases the noise becomes unacceptably high at high iteration numbers. Stopping iterations early is dangerous: because the convergence of ML-EM depends on position and even orientation, the spatial resolution can be very position dependent at low iteration numbers. This is illustrated by the PET simulation in Figure 2. The true object is a circular ring of activity, embedded in a large disc with background activity. Attenuation within the disc is uniform. Because the attenuation along central projection lines is higher, the convergence of ML-EM is slower in the centre. At 10 iterations, the spatial resolution in the centre is clearly lower than near the edge. At 100 iterations, the resolution has become uniform. However, when Poisson noise is present, the noise level at 100 iterations is very high.

6.1 MAP with Gibbs Prior

As discussed above, maximising the likelihood is equivalent to maximising the posterior, if one can assume that nothing is known about the image, prior to the measurement. But obviously, some a-priori knowledge is available: the image should not be too noisy. Somehow, this prior knowledge must be translated into a mathematical expression, which can be inserted in equation (4).

A convenient way to define a prior distribution favouring smooth solutions is via a Markov random field (or equivalently: a Gibbs random field). In a Markov random field, the probability of a particular voxel depends on the voxel values in a neighbourhood of that voxel. The dependence is defined with a Gibbs distribution of the following form (e.g. Geman and McLure[22]):

$$\text{prob}\left[f_j\right] = \frac{1}{Z} e^{-\beta U(f)} \text{ with } U(f) = \sum_{k \in N_j} V\left(f_j, f_k\right) \tag{17}$$

where prob[f_j] denotes the conditional probability for the value in voxel j given the rest of the image, Z is a normalization constant, β a factor that determines the strength of the prior (1/β is the "temperature"), and N_j contains the voxels neighbouring voxel j. U is the so-called energy function, and higher energies are less likely. In fact, more complex forms of U are allowed, but this one is nearly always used. To favour smooth images, noisy ones should be penalized with a higher energy (hence the alternative name "penalized likelihood"). An obvious choice for V is the "quadratic penalty":

$$V_Q(f_j, f_k) = \frac{(f_j - f_k)^2}{2\sigma^2} \tag{18}$$

which makes *prob[f]* maximum for a perfectly uniform image. A possible disadvantage of V_Q is that it heavily penalizes large differences, causing strong smoothing over edges. To avoid that, the Huber function (e.g. Mumcuoğlu *et al.*[23]) quadratically penalizes "small" differences, but penalizes "large" differences only linearly:

$$\begin{aligned} V_H(f_j, f_k) &= \frac{(f_j - f_k)^2}{2\sigma^2} && \text{if} |f_j - f_k| < \sigma \\ &= \frac{|f_j - f_k| - \sigma/2}{\sigma} && \text{if } |f_j - f_k| \geq \sigma \end{aligned} \tag{19}$$

A function with a similar edge-preserving characteristic (but with a gradual transition to the linear behaviour) is the logcosh function (see e.g. de Pierro[24]): $V_L = \ln(\cosh((f_j - f_k)/\sigma))$. Functions have been devised that preserve edges better than V_H or V_L, e.g. by applying a constant instead of a linear penalty for large differences. However, in contrast to the functions mentioned above, functions applying a constant penalty are not concave; they introduce multiple local maxima, making the final MAP-reconstruction dependent upon the initial image and the particular optimisation algorithm. For that reason, these functions are rarely used in emission tomography.

The functions V_Q, V_H and V_L penalize in proportion to the (square of) absolute differences between neighbouring voxels. But because the absolute tracer uptake values vary with scan time, injected dose, attenuation, patient weight and metabolism, it can be difficult to select a good value for the parameter σ. As one can see from the equations, ML-EM is insensitive to the absolute counts in the sinogram: if the sinogram is multiplied with a factor, the corresponding ML-EM reconstruction will simply scale with the same factor. This is no longer true for the MAP-reconstruction based on the quadratic or Huber prior. To avoid this problem, a concave function penalizing relative differences rather than absolute ones has been proposed:[25]

$$V_R(f_j, f_k) = \frac{(f_j - f_k)^2}{f_j + f_k + \gamma|f_j - f_k|} \tag{20}$$

Similar to the Huber prior, V_R has some edge-preserving features: the penalty decreases slower when the relative difference $|f_j - f_k|/(f_j + f_k)$ is large compared to $1/\gamma$.

Having defined the prior, an algorithm is needed to maximise the posterior, or equivalently, its logarithm: $\ln \text{prob}[P|f] + \ln \text{prob}[f]$. Direct application of the EM strategy yields the following expression:

$$f_j^{new} = \frac{f_j^{old}}{\sum_l a_{lj}} \sum_i a_{ij} \frac{p_i}{\sum_k a_{ik} f_k^{old} + \beta \frac{\partial U(f)}{\partial f_j}} \tag{21}$$

where $U(f)$ is the energy function of the Gibbs prior. The problem is that the derivative must be evaluated in f_j^{new}, which complicates the solution dramatically. Green[26] proposed the one step late (OSL) approach, which is to evaluate the derivative in the available reconstruction image f_j^{old}. This approach is effective if β is not too large, but convergence is not guaranteed, and with large β the algorithm can actually diverge. de Pierro[24] found a new (non-statistical) derivation for the classical ML-EM algorithm, which is based on the construction of a surrogate function at each iteration. This is a function that coincides with the likelihood function up to the first derivate and which is easier to maximise than the likelihood. Each iteration, the surrogate is maximised, and one can show that this guarantees an increase of the likelihood. The same approach can be applied to most Gibbs priors, and hence, to the posterior. This has led to the development of new and more robust algorithms for MAP-reconstruction (see Ahn and Fessler[27] and the references therein).

6.2 *Post-smoothing Unconstrained ML-EM*

As shown in Figure 2, possible convergence problems of ML-EM are eliminated by iterating sufficiently long. This leads to unacceptable noise levels, but the noise can be reduced considerably with moderate smoothing,[21,28] as illustrated in Figure 3. Since this is high frequency noise, smoothing is even more effective than with uncorrelated noise.[29]

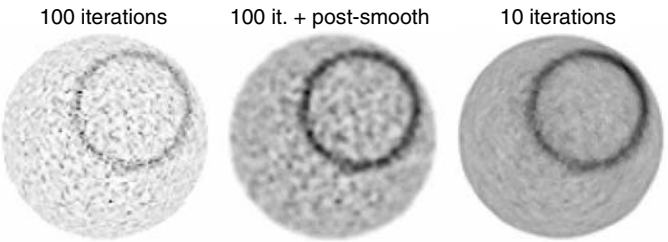

FIGURE 3. Same PET simulation as in Figure 2, comparing post-smoothing ML-EM after 100 iterations with the reconstruction without post-smoothing (left) and the ML-EM reconstruction at 10 iterations.

An advantage of post-smoothed ML-EM is its wide availability. It is currently available on most SPECT and PET systems. As will be discussed below, an additional advantage (for some applications) is that it can impose nearly uniform spatial resolution. A disadvantage is that many iterations are required to achieve this resolution uniformity.

6.3 Expectation-Maximisation-Smooth (EMS)

Some groups have investigated the application of a smoothing filter after every ML-EM iteration, called EMS (expectation-maximisation-smooth) or inter-iteration filtering.[28,30,31] Although there is no proof of convergence, this algorithm always appears to converge in practice. The image characteristics are somewhat similar to those of penalized likelihood reconstruction.

6.4 Uniform Resolution

For some applications in emission tomography it is desirable that the spatial resolution in the images is predictable, independent of the object and usually also independent of position. Examples are quantitative analysis of PET studies, where one has to assume that changes in numbers are due to changes in metabolism, and not to changes in characteristics of the imaging system.

A problem with most MAP-algorithms (and also with EMS) is that they tend to produce object and position dependent resolution. The likelihood term seeks ideal resolution, while the prior (penalty) wants no resolution. The final resolution depends on the relative strength of both terms. The strength of the prior is usually chosen to be constant, but the strength of the likelihood depends on the object, on the position within the object and even on the orientation. To obtain the same balance between the two everywhere, the strength of the prior has to follow that of the likelihood. Such algorithms have been designed, and compared to post-smoothed ML-EM.[32-34] It is found that at matched resolution, the noise characteristics of these penalized- likelihood algorithms are similar to those of post-smoothed ML-EM. Post-smoothed ML-EM is much easier to program, and is already widely available. However, the penalty improves the conditioning of the inverse problem, which can be exploited to obtain faster convergence.[34] So for applications requiring uniform resolution, penalized-likelihood may still have its role as an acceleration technique.

6.5 Anatomical Priors

Many attempts have been undertaken to exploit anatomical knowledge, available from registered MRI or CT images, during the reconstruction of PET or SPECT data. The aim is to avoid the resolution loss due to the regularization, or even to recover the resolution of the emission reconstruction, by making use of the superior resolution of the anatomical images.

These attempts are only meaningful if one can assume that there is a fair overlap between anatomical and functional (tracer uptake) boundaries. Fortunately, this seems to be a valid assumption in many applications.

The anatomical information has been used to tune the noise suppressing prior in a MAP-algorithm, by limiting smoothing to within organ boundaries revealed by the anatomical data.[35-40] In addition, the segmented MR image can be used for attenuation and scatter correction purposes.[41] If the finite resolution of the emission tomograph is modelled, then these algorithms can produce a strong resolution recovery near anatomical boundaries. To further push resolution recovery, Sastry and Carson[42] introduced a tissue composition model, which considers each (coarse) PET-voxel as composed of one or more tissue classes, which are obtained from a segmented MRI image of the same patient. The PET-activity in every voxel is then a weighted sum of tissue class activities. Figure 4 shows an example obtained with a similar approach.

6.6 *Median Root Prior*

Alenius *et al.*[45,46] have proposed a penalized-likelihood method based on a median filter. The algorithm can be written in the following form:

$$f_j^{new} = \frac{f_j^{old}}{\sum_i a_{ij}} \sum_i a_{ij} \frac{p_i}{\sum_k a_{ik} f_k^{old} + \beta \frac{f_j^{old} - M_j}{M_j}} \tag{22}$$

where M_j is the value of voxel j, obtained by median filtering the image f_j^{old}. Following the OSL interpretation discussed above, the difference between a voxel and the median of its neighbourhood is used as the gradient of some energy function. The derivation is empirical, and in fact, the corresponding energy function does not exist. Hsiao[47] recently proposed a very similar

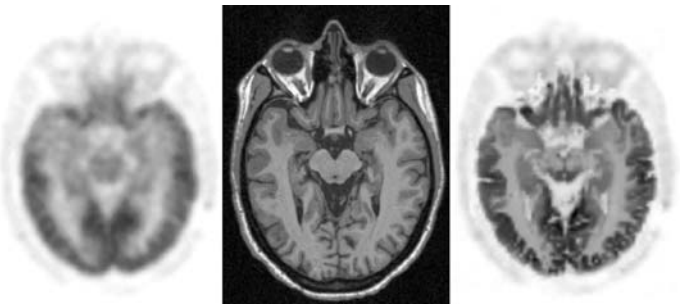

FIGURE 4. PET image obtained with MAP-reconstruction without (left) and with (right) the use of anatomical information from a registered MRI image (center), using a tissue composition model.[43] The 3D sinogram was first rebinned with FORE,[44] then the MAP-reconstructions were computed.[12]

algorithm, which does minimize a well defined objective function. Albeit empirical, the algorithm is very effective and has some interesting features. In contrast to most priors, it does not strive towards a completely flat image (which is not a very realistic prior assumption). Instead, it only requires that the image be locally monotonic. Locally monotonic images do not have small hot spots, but they can have sharp and/or smooth edges. In addition, similar to the relative difference prior, it penalizes relative differences, making β a "unit-less", easily tuned parameter. Stated intuitively, the MRP algorithm essentially suppresses all hot or cold spots that are small compared to the size of its median filter. It follows that, when applied in hot spot detection applications (such as PET whole body imaging in oncology), the mask should be chosen sufficiently small and β not too large.

Figure 5 illustrates the behaviour of ML-EM with moderate post-smoothing, MRP and the relative difference prior. Large homogenous regions (such as the liver and the mediastinum) are much better visualized in the MAP-reconstruction images, in particular with MRP. Smaller structures, such as the blood vessels, are somewhat attenuated by both priors. A very hot and small spot, such as the lesion in the neck, is strongly suppressed by the MRP-penalty. It is better preserved by the relative difference prior, because of its tolerance for large voxel differences. But it is best recovered by the noisier ML-EM image,

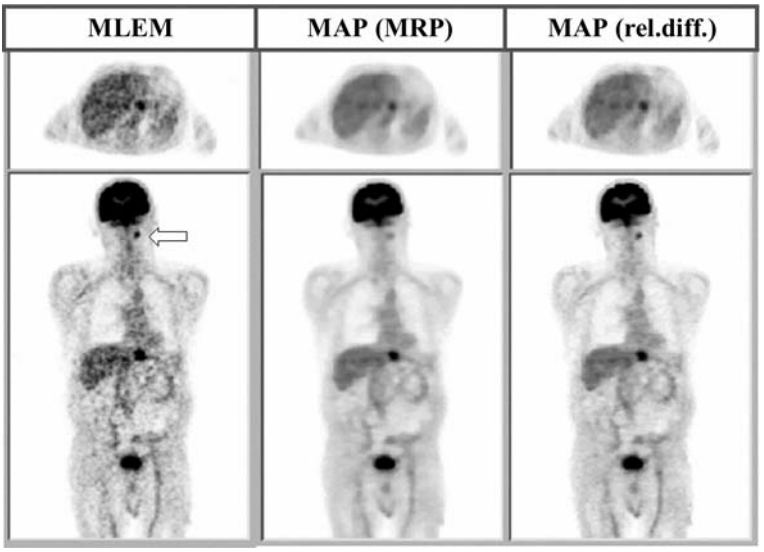

FIGURE 5. Three reconstructions of the same patient's PET scan. Left: post-smoothed ML-EM; center: MAP with median root prior; and right: MAP with relative difference prior (V_R). The median root prior has superior noise suppression and edge preservation for larger structures. However, hot spots that are small compared to the median filter size are suppressed (see the lesion in the neck, indicated by the arrow).

because only moderate smoothing was applied. The penalties should be applied with caution and must be optimized for each particular task.

7. Approaches to Acceleration of ML-EM

7.1 Ordered Subsets EM: OS-EM

The computation time for iterative reconstruction algorithms is typically orders of magnitude longer than for FBP. Fortunately steps can be taken to significantly accelerate these algorithms. The most widely used acceleration technique, perhaps due to its simplicity, is the ordered subsets EM or OS-EM algorithm.[48] Ordered subsets or block-iterative methods have been applied to a number of iterative algorithms with OS-EM being similar in some respects to MART. The essential difference between OS-EM and ML-EM is the use of only a subset of projections for updating rather than comparison of all estimated and measured projections. For OS-EM one iteration is normally considered the use of all data once; consequently use of only part of the data during the update process is termed a sub-iteration. In OS-EM the reconstruction proceeds by utilising subsets of the projections, chosen in a specific order that attempts to maximise the new information being added in sub-iterations. Iteration proceeds by using different projections in each subsequent subset until all projections are used. The resulting equation is very similar to the standard ML-EM equation, the only difference being the use of subsets, S_n, where $n \in N$, the total number of projections divided by the number of projections per subset or subset size:

$$f_j^{new} = \frac{f_j^{old}}{\sum\limits_{i \in S_n} a_{ij}} \sum_{i \in S_n} a_{ij} \frac{p_i}{\sum\limits_k a_{ik} f_k^{old}} \tag{23}$$

Note that sub-iterations recycle through S_n for subsequent iterations. Provided subset size is not too small and subset order is chosen carefully, the reconstruction estimate of OS-EM at each sub-iteration is almost indistinguishable from the reconstruction for a corresponding number of full iterations of ML-EM (Figure 3). The computation time for each sub-iteration is roughly equal to the time for a conventional ML-EM iteration divided by N, due to the reduced number of forward and back projection operations in the subset. Consequently for 128 projections and subset size of 4, the acceleration factor is approximately 32.

There are some theoretical concerns regarding OS-EM. For example, there is no proof that OS-EM in its standard form converges to the same solution as ML-EM. In fact OS-EM reaches a limit cycle, where the result depends to a small extent on the point in the subset sequence at which iteration is stopped. However, with the noisy data typical of clinical practice, this has not been found to be of particular concern, since the possible

solutions tend to be very similar, provided subset size is not too small. Choosing subset order is important as a necessary condition is that there be subset balance, which requires that $\sum_{i \in S_n} a_{ij}$ is independent of n. In practice the use of four well spaced projections is a reasonable compromise between speed and reconstruction quality.[49]

7.2 Variants of OS-EM

There have recently been a number of variants of OS-EM that address some of its theoretical limitations. The rescaled block-iterative (RBI) algorithm[50] does converge to a single solution in the absence of noise, independent of subset order. The equation for RBI is

$$f_j^{new} = f_j^{old} + \frac{f_j^{old}}{\max \sum_{i \in S_n} a_{ij}} \sum_{i \in S_n} a_{ij} \left(\frac{p_i}{\sum_k a_{ik} f_k^{old}} - 1 \right) \tag{24}$$

which can be seen to reduce to exactly OS-EM when $\sum_{i \in S_n} a_{ij}$ is constant for all n. Alternatively the row-action maximum likelihood algorithm (RAMLA) is designed to avoid the limit cycle by using under-relaxation (deliberately decreasing the influence of the update as iteration advances).[51] The equation again has a very similar form:

$$f_j^{new} = f_j^{old} + \lambda_k f_j^{old} \sum_{i \in S_n} a_{ij} \left(\frac{p_i}{\sum_k a_{ik} f_k^{old}} - 1 \right), \quad 0 < \lambda_k \leq \sum_{i \in S_n} a_{ij}^{-1} \tag{25}$$

where λ_k decreases with iteration k, but remains constant throughout the sub-iterations within each main iteration. More recent work suggests that λ could be varied with sub-iteration.[52] An alternative approach to controlling the limit cycle is to increase the subset size with iteration,[48] providing results similar to over-relaxation.[53] The similarity in these algorithms leads to some confusion in their applicability, particularly as some of the theoretical concerns are of little clinical relevance e.g. with noisy data the ML solution is undesirable. Nevertheless, as seen from the above equations, some level of standardisation in the approaches is emerging.

7.3 Acceleration for Regularized Reconstruction Algorithms

The same approaches to acceleration can be applied to regularised statistical algorithms as to ML-EM since their formulations are generally similar. The OS-GP was suggested as the regularised OS-EM equivalent for MAP-OSL reconstruction based on the Gibbs prior.[48] Similarly block-iterative approaches have been used for the basis of developing the RBI-MAP

algorithm[54] and the block sequential regularized EM algorithm (BSREM) as a regularized form of RAMLA.[55] There is some debate regarding the choice of regularized acceleration algorithm, with suggestion that BSREM type algorithms have theoretical advantages.[27,55] Still further work is necessary to confirm which algorithm is optimal for clinical application.[56]

8. Using ML with Non-Poisson Data

8.1 Transmission

The ML-EM algorithm described above is based on the assumption that the data can be considered as linear combinations of the unknown tracer distribution, and that they are Poisson distributed. In transmission tomography, the counts t_i are Poisson distributed, but they are a non-linear function of the image μ_j. For that purpose, dedicated ML and MAP algorithms have been developed for transmission tomography. The likelihood function to be optimised is now:

$$L_t = \sum_i t_i \ln \hat{t}_i - \hat{t}_i, \text{ with } \hat{t}_i = b_i \exp\left(-\sum_j l_{ij}\mu_j\right) + s_i \tag{26}$$

where t_i is the measured transmission count rate, b_i is the blank scan, l_{ij} is the intersection length of projection line i with the image voxel j, and s_i contains additive contributions, such as Compton scatter, random coincidences or emission activity in post- injection transmission scanning in PET.

One can apply the expectation-maximisation (EM) strategy again.[57,58] However, in this case, it leads to a somewhat cumbersome expression. To obtain more elegant algorithms, other optimisation methods have been used. Mumcuoğlu *et al.*[59] proposed a conjugate gradient method. Using de Pierro's surrogate function approach, Fessler[60] derived a suitable class of algorithms. The unconstrained and non-parallel version can be written as:

$$\mu_j^{new} = \mu_j^{old} + \frac{\sum_i l_{ij}(1 - t_i/\hat{t}_i)(\hat{t}_i - s_i)}{d_j} \text{ with } d_j = \sum_i \frac{l_{ij}^2 (t_i - s_i)^2}{\alpha_{ij} t_i} \tag{27}$$

where α_{ij} can be chosen, as long as $\alpha_{ij} \geq 0$ and $\sum_j \alpha_{ij} > 0$.

Setting $\alpha_{ij} = l_{ij}\mu_j^{old} / \sum_k l_{ik}\mu_k^{old}$ yields the convex algorithm of Lange and Fessler,[18] choosing $\alpha_{ij} = l_{ij} / \sum_k l_{ik}$ produces the algorithm used in Nuyts.[61] These algorithms behave similarly to the ML-EM algorithm in emission tomography: they have similar noise propagation; they need a comparable amount of computation time and show similar convergence speed. They can be accelerated with ordered subsets as well.[62]

As in emission tomography, the noise can be suppressed by combining the likelihood with a prior. However, in transmission tomography more prior

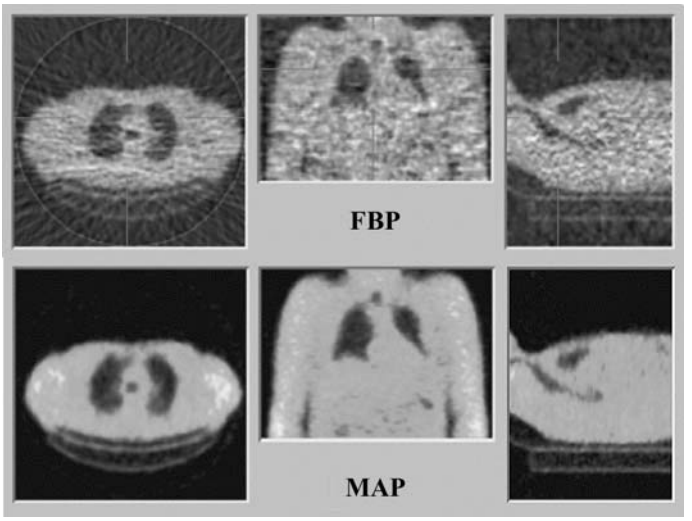

FIGURE 6. Comparison of FBP (above) and MAP (below) reconstruction of very short PET transmission scan.

information is available: the attenuation of tissue and bone is known. In addition, attenuation images are much smoother than emission images. For this reason, more aggressive priors can be applied.[59-61] An example is shown in Figure 6. The effect of these strong priors can be considered as a segmentation, thus producing transmission maps resembling those of the so-called "segmented attenuation correction" algorithms.[63-66] The main difference with the MAP-approach is that the segmentation is done during, not after the reconstruction.

8.2 Corrected Emission

To reduce processing time and data storage, it is often convenient to work with pre-corrected data. For example, careful reconstruction on a PET system requires separate storage of at least three sinograms: one containing the raw, Poisson distributed data, one containing multiplicative effects (including attenuation and detector efficiency) and another one containing additive effects (Compton scatter and randoms). In particular in 3D PET, this represents an impressive amount of data, requiring considerable processing time for reconstruction. An attractive alternative is to produce a single smaller sinogram by pre-correcting the data and applying Fourier rebinning. But then the data are no longer Poisson distributed. It has been shown that ML-EM yields suboptimal images from such data.[12]

An obvious solution is to abandon ML-EM and use a weighted least squares algorithm instead (e.g. Fessler[11]). However, because the ML-EM algorithm is more widely available, an alternative solution has been devised.

It is possible to manipulate the data in order to restore the Poisson distribution approximately, and then undo the manipulation during ML-EM reconstruction. Consider a random number x, and assume that estimates for its mean and variance are available. The distribution can be shifted, such as to make the variance equal to the mean, as required by a Poisson distribution: mean $(x + s) = (x + s)$. It follows that the shift amount to $s = (x) - \text{mean}(x)$. This approach, for example, is used to deal with randoms correction in PET.[67] Alternatively, one can scale x, such that mean$(a\ x) =$ var$(a\ x)$, which yields $a = \text{mean}(x)\ /\ \text{var}(x)$. This approach can be used for reconstruction of transmission scans, if no dedicated transmission algorithm is available.[68]

9. Application to Specific Problems

9.1 More Complex Emission Models

The attraction of iterative reconstruction is that it can be relatively easily modified so as to be applicable to specific emission models. The limitation tends to be that the computational speed is directly related to the complexity of the underlying model. Not only does the computation per iteration increase but also the number of iterations required. Fortunately the combination of faster computers and effective acceleration models means that tackling more complex problems can now be contemplated.

The incorporation of information on distance-dependent resolution in the transition matrix[69,70] results in the iterative reconstruction effectively operating in 3D, allowing for the probability that photons travelling obliquely to the detector still can be detected. Incorporation of resolution loss in the system model in theory suggests that the final reconstruction should have resolution recovery, although in practice this is limited by slow convergence. This will be discussed in more detail elsewhere. The correction of scatter is a particularly demanding problem due to the complex dependence of scatter on the non-homogeneous attenuation in tissue and the broad 3D distribution of scatter. Modelling scatter for incorporation in the transition matrix is difficult since it is object dependent, although appropriate models have been developed[71,72] and efficient Monte Carlo solutions have been proposed.[73] It is worthwhile noting that there are advantages to incorporating measured scatter estimates (such as might be obtained from multiple energy windows) directly in the projector.[74] The incorporation of estimated scatter ($\hat{s}$) in the ML-EM equation is given by:

$$f_j^{new} = \frac{f_j^{old}}{\sum_i a_{ij}} \sum_i a_{ij} \frac{p_i}{\sum_k a_{ik} f_k^{old} + \hat{s}} \tag{28}$$

Since the scatter estimate is simply used to provide consistency between estimated and measured projections the resultant reconstruction is much more resistant to noise than if the scatter was directly subtracted from the measured projections. Indeed the direct subtraction introduces concern that the projections are no longer Poisson distributed (as discussed further below). A similar approach can be taken to account for measured randoms in PET or indeed any source of measured counts that is not accounted for by the transition matrix.

Extensions to the system modelling have been developed to accommodate the more complex situation with either dual radionuclides[75] or multiple-energy radionuclides.[76] In dual radionuclide studies two distributions are simultaneously determined allowing for cross-talk between the radionuclides being imaged. In the case of a multiple-energy radionuclide a single estimate can be determined taking into account the different attenuation that occurs at different emission energies.

A further complexity to iterative reconstruction can be introduced if a fourth dimension is included in the model to account for tracer redistribution with time or motion occurring during gated acquisition. Provided a model of the time variation is well understood this can be incorporated in the transition matrix so as to account for inconsistencies that may otherwise occur between estimated and measured projections. The interested reader is referred to relevant publications.[77,78]

9.2 3D Reconstruction

Some of the complexities mentioned above already suggest the need for a 3D reconstruction (or even in some cases 4D) in order to cater for the possibility of non-planar detection. Clearly this becomes even more evident with detectors that specifically are designed to acquire data in 3D (e.g. cone-beam SPECT, pinhole SPECT or septa-less PET). The attraction of iterative reconstruction is that the basic algorithms do not require any specific change except for definition of the transition matrix, unlike FBP, which usually requires the derivation of specific non-standard filters. A specific example where iterative reconstruction has been widely applied is in 3D PET where usually some form of rebinning is first performed (e.g. Fourier rebinning or FORE) in combination with iterative reconstruction.[44] Alternatives based on direct 3D iterative reconstruction have also been implemented.[79,80]

9.3 Motion Correction

An area where iterative reconstruction is proving particularly useful is the correction of motion, determined either from independent measurement[81] or via estimation directly from the measured projections.[82] Since patient movement can be considered as equivalent to detector movement, the reconstruction necessitates the incorporation of data at specific angular positions

relative to the reconstructed object space. Provided most projections can be considered to correspond to a stationary patient at various locations, reconstruction can be accomplished rather easily using iterative reconstruction with OS-EM selecting subsets of projections so as to correspond to specific patient/detector positions. More exact correction necessitates either multiple short-duration acquisitions to minimise the possibility of movement in an individual acquisition frame or use of list-mode[83] where each individual event can be corrected for movement so as to locate the exact ray-path. The EM list mode equation[84,85] can be formulated to be very similar to the conventional EM equation:

$$f_j^{new} = \frac{f_j^{old}}{\sum_i a_{ij}} \sum_{i \in L} a_{ij} \frac{1}{\sum_k a_{ik} f_k^{old}} \tag{29}$$

Note, however, that the main summation is restricted to the elements in the list-mode data rather than all projection bins; this effectively ignores zero projection bins and consequently the numerator reduces to unity. The OS-EM version simply divides the list-mode data into sub-lists.

10. Clinical Evaluation

Evaluation and clinical validation of image reconstruction algorithms is inherently difficult and sometimes unconvincing. There is a clear need for guidelines to evaluate reconstruction techniques and other image processing issues in emission tomography. A particular concern in clinical studies is the tendency to compare not only different algorithms but different approaches to processing, without effort to isolate the effects due to the reconstruction algorithm itself. Examples are the comparison of iterative algorithms including attenuation correction with analytic approaches without attenuation correction or the comparison of different iterative algorithms where the fundamental implementation differs (e.g. use of "blobs"[86] rather than pixels in the projector). This simply adds to the confusion in interpreting results. A further common problem is the comparison of clinical images where the reconstruction algorithm results in different signal to noise properties, typically dependent on the number of iterations utilised. Evaluation for a range of parameters tends to provide more objective results where the trade-off in noise and signal (e.g. recovery coefficient) can be more meaningfully compared.

Most of the algorithms developed so far have been evaluated using either simulated or experimentally measured phantom studies, in addition to qualitative evaluation of clinical data. This has been extended more recently to objective assessment of image quality using Receiver Operating Characteristics (ROC) analysis based on human or computer observers,[87] evaluation of the influence of reconstruction techniques on tracer kinetic

parameter estimation[88] and voxel-based analysis in functional brain imaging using statistical parametric mapping.[89,90]

10.1 Validation Using Simulation and Phantom Studies

A simulator is an efficient tool that can be used to generate data sets in a controllable manner in order to assess the performance of different reconstruction algorithms. Medical imaging simulation tools have been shown to be very useful for validation and comparative evaluation of image reconstruction techniques since it is possible to obtain a reference image to which reconstructed images should be compared. The ability to theoretically model the propagation of photon noise through emission tomography reconstruction algorithms is crucial in evaluating the reconstructed image quality as a function of parameters of the algorithm. Two broad categories of simulation packages have emerged: simplified analytical modelling tools and sophisticated Monte Carlo simulations. In the first class of simulators, several simplifying approximations are adopted to improve ease of use and speed of operation (e.g. Gaussian noise distribution and scatter model based on an analytical point spread function).[91] On the other hand, the Monte Carlo method is widely used for solving problems involving statistical processes and is very useful in medical imaging due to the stochastic nature of radiation emission, transport and detection processes. Many general purpose and dedicated Monte Carlo programs have been in use in the field of nuclear imaging with many of them available in the public domain.[92,93] Although variance reduction techniques have been developed to reduce computation time, the main drawback of the Monte Carlo method is that it is extremely time-consuming.

Software and physical phantoms used in medical imaging were historically limited to simple point, rod, and slab shapes of sources and attenuating media. Such simple geometries are useful in studying fundamental issues of image reconstruction, but clinically realistic distributions cannot be evaluated by such simple geometries. A precise modeling of the human body requires appropriate information on the location, shape, density, and elemental composition of the organs or tissues. Several software phantoms modeling different parts of the human body have been developed over the years to assess the accuracy of reconstruction procedures and are described in detail in Chapter 11 of this book. Interested readers are also refereed to textbooks referenced above discussing extensively issues related to Monte Carlo modeling in nuclear medicine.

10.2 Subjective and Objective Assessment of Image Quality

Different approaches have been suggested to judge image quality when evaluating image reconstruction algorithms. As the 'best' reconstruction algorithm can only be selected with respect to a certain task, different

'basic' performance measures can be used. Basically, there is no single figure of merit that summarises algorithm performance, since performance ultimately depends on the diagnostic task being performed. Well-established figures of merit known to have a large influence on many types of task performance are generally used to assess image quality.[94]

In a clinical environment, the evaluation is further hampered by the multiplicity of the medical purposes for which the corrections may be studied. In the simplest approach, trained nuclear medicine physicians carry out observer performance studies and are asked to rank images by their degree of quality. A common method to assess image quality with respect to a detection task is the use of observer studies where the performance of any observer (human or computer algorithm) is characterized by ROC analysis. In such studies, observers rate images based on their confidence that a defect/lesion exists in a large set of images. Curve-fitting methods are then used to fit the rating data to receiver operating characteristic (ROC) curves, which plot the covariation in ''true positive'' and ''false positive'' conditional probabilities across changes in the decision-variable's criterion for a ''positive'' binary decision.[87,95,96] The estimated area under the fitted ROC curve is often used as a general index of image quality or performance accuracy for any alternative classification task. ROC and localization ROC (LROC) techniques have been extensively used to evaluate lesion detectability.[97] For example, using simulated MCAT phantom data and randomly located 1 cm-diameter lesions, Lartizien *et al.*[98] demonstrated that the FORE+AWOSEM (attenuation weighted OS-EM) algorithm results in the best overall detection and localization performance for 1-cm-diameter lesions compared with the FORE+OSEM and FORE+FBP algorithms in PET imaging. The major drawback of this approach is that it is costly and complex, since a reasonable number of experienced observers should be used to analyse many images under carefully controlled conditions. In addition, such techniques rely on experimental phantom measurements or simulated data since the ground truth needs to be known. Furthermore, for optimisation of reconstruction algorithms in which possible parameter settings suffer a combinatorial explosion, human psychophysical studies are simply not viable. Therefore, most qualitative assessment studies are restricted to subjective observer rating scores by experienced physicians. The diversity of available algorithms also makes comparison difficult as results can be inconclusive when limited cross-comparison is involved.

There are an increasing number of published articles that verify the benefit of utilising iterative algorithms in clinical situations. These verify many of the theoretical advantages outlined in section 2.2, including ease of incorporating attenuation correction,[99] reduction in streak artefacts,[100] tolerance to missing data[101] and most importantly, noise reduction in low count areas of the image. It should be noted that, since noise generated in iterative reconstruction is signal-dependent, the signal to noise gains are restricted

to low count areas, whereas high count areas may at best be unaffected.[102,103] In many nuclear medicine studies there is interest in derived parameters rather than individual images (e.g. extracted functional parameters or significance maps reflecting changes in activity distribution). It therefore is of some interest to identify which reconstruction algorithms are optimal for these applications and to verify that the reconstruction algorithms provide quantitative results.[104,105] An interesting approach in comparative evaluation studies for functional brain imaging is to carry out voxel-based statistical analysis using statistical parametric mapping (SPM).[106] A recent study on the impact of model-based scatter correction and iterative reconstruction on spatial distribution of ^{18}F-[FDG] in reconstructed brain PET images of healthy subjects using this kind of analysis demonstrated that OS-EM reconstruction does not result in significant changes when compared to FBP reconstruction procedures, while significant differences in ^{18}F-[FDG] distribution arise when images are reconstructed with and without explicit scatter correction for some cerebral areas.[89] Other related studies showed that iterative reconstruction has the potential to increase the statistical power and to give the best trade-off between signal detection and noise reduction in PET activation studies as compared with FBP reconstruction.[90]

11. Future Prospects

The progress in iterative reconstruction has been immense in the past ten years, the main opportunities arising from the availability of both improved processing speed and faster algorithms. This has permitted much more ambitious algorithms that tackle not just conventional 2D reconstruction but a range of applications in 3D (e.g. Liu *et al.*[80]) and even 4D (including the time domain).[78] The appeal of iterative reconstruction in adapting to different acquisition geometry or physical situation has already been well demonstrated. The practicality of performing such reconstruction is revitalising the consideration of alternative approaches to imaging such as use of multiple pinhole collimators or so-called 'Compton' imaging (e.g. Braem *et al.*[107]). Although currently focussed on high resolution imaging for small animals there is potential for future developments to address specific clinical applications.

Anticipating even greater gains in computer power, even more complex and ambitious reconstruction becomes clinically realistic. More extensive use of fully 3D reconstruction algorithms can be expected, with more exact correction likely to provide optimal quality of reconstruction. The feasibility of using Monte Carlo-based approaches has been demonstrated[108] and previous work has demonstrated that symmetries can be used to greatly accelerate the Monte Carlo calculations.[73] Whether this degree of sophistication in accurately modelling the emission and detection processes is warranted in clinical practice remains to be seen.

The demands placed on reconstruction algorithms continue to expand. There is a strong evolution towards more complex acquisition hardware, in particular in PET (more and smaller detector crystals, larger axial extent of the detectors, better energy resolution allowing multi-energy acquisitions). In some of these complex geometries, the assumptions of Fourier rebinning are no longer valid and more accurate rebinning algorithms are needed.[44] The huge size of the data sets compensates for the growing computer speed and, since acceleration of 3D reconstruction has probably reached its limits, the computational demand continues to push technology to its limit.

The availability of dual-modality instruments is likely to influence the approaches to reconstruction, with algorithms that combine anatomical information likely to increase in popularity. These approaches have previously relied on accurate registration and ready access to anatomical data. The advent of dual modality instruments makes this much more practical.

Accompanying these developments is the increasingly difficulty task of understanding the intricacies of the ever-increasing range of algorithms. In the clinical environment the uncertainty in choice of algorithm has acted as a deterrent to their use. To some extent there is 'convergence' in the various algorithms and it is to be anticipated that there may be some future consensus as to the best algorithm to choose for specific applications. The availability of open platforms and more widely available software, specifically designed for clinical application, should at some stage accelerate the acceptance of algorithms. However, at the time of writing OS-EM remains the most widely used iterative algorithm in clinical practice, with variants specifically tailored for use with PET or transmission reconstruction. It is clear, however, that there is great need for development of standard approaches for evaluation of algorithms e.g. the availability of standard validated data for inter-lab comparison.

Acknowledgments. One of the co-authors (HZ) acknowledges the support of Grant SNSF 3152A0-102143 awarded by the Swiss National Science Foundation.

References

1. Lalush D. S., "Iterative image reconstruction" in: *Emission Tomography: The Fundamentals of PET and SPECT*. edited by M Wernick and J Aarsvold Academic Press, San Diego, (2004) *in press*
2. Fessler J. A., "Statistical image reconstruction methods for transmission tomography" in: *Handbook of Medical Imaging*, edited by M. Sonka and J. M. Fitzpatrick SPIE, Bellingham, (2000) *pp* 11-70.
3. Lewitt R. M. and Matej, S., Overview of methods for image reconstruction from projections in emission computed tomography. *Proceedings of the IEEE* **91:** 1588-1611 (2003).

4. Herman G. T. and Meyer, L. B., Algebraic reconstruction techniques can be made computationally efficient. *IEEE Trans Med Imaging* **12:** 600-609 (1993).
5. Bender R., Gordon, R. and Herman, G. T., Algebraic reconstruction techniques (ART) for three-dimensional electron microscopy and x-ray photography. *J Theor Biol* **29:** 471-482 (1970).
6. Lu H. H., Chen, C. M. and Yang, I. H., Cross-reference weighted least square estimates for positron emission tomography. *IEEE Trans Med Imaging* **17:** 11-8 (1998).
7. Guan H., Gordon, R. and Zhu, Y., Combining various projection access schemes with the algebraic reconstruction technique for low-contrast detection in computed tomography. *Phys Med Biol* **43:** 2413-2421 (1998).
8. Byrne C. L., Block-iterative methods for image reconstruction from projections. *IEEE Trans Image Proc* **5:** 792-794 (1996).
9. Barrett H. H., Wilson, D. W. and Tsui, B. M. W., Noise properties of the EM algorithm. I. Theory. *Phys Med Biol* **39:** 833-846 (1994).
10. Sauer K. and Bouman, C., A local update strategy for iterative reconstruction from projections. *IEEE Trans Sign Proc* **41:** 534-548 (1993).
11. Fessler J. A., Penalized weighted least-squares image reconstruction for positron emission tomography. *IEEE Trans Med Imaging* **13:** 290290-300 (1994).
12. Comtat C., Kinahan, P. E., Defrise, M. *et al*., Fast reconstruction of 3D PET data with accurate statistical modeling. *IEEE Trans Nucl Sci* **45:** 1083-1089 (1998).
13. Meikle S. R., Fulton, R. R., Eberl, S. *et al*., An investigation of coded aperture imaging for small animal SPECT. *IEEE Trans Nucl Sci* **48:** 816-821 (2001).
14. Kaufman L., Maximum likelihood, least squares, and penalized least squares for PET. *IEEE Trans Med Imaging* **12:** 200-214 (1993).
15. Chinn G. and Huang, S. C., A general class of preconditioners for statistical iterative reconstruction of emission computed tomography. *IEEE Trans Med Imaging* **16:** 11-10 (1997).
16. Dempster A. P., Laird, N. M. and Rubin, D. B., Maximum likelihood from incomplete data via the EM algorithm. *J Royal Stat Soc B* **38:** 11-38 (1977).
17. Shepp L. A. and Vardi, Y., Maximum likelihood reconstruction for emission tomography. *IEEE Trans Med Imag* **1:** 113-122 (1982).
18. Lange K. and Fessler, J. A., Globally convergent algorithms for maximum a posteriori transmission tomography. *IEEE Trans Imag Proc* **4:** 1430-1438 (1995).
19. Miller M. I., Snyder, D. L. and Miller, T. R., Maximum likelihood reconstruction for single-photon emission computed tomography. *IEEE Trans Nucl Sci* **32:** 769769- 778 (1985).
20. Wilson D. W., Tsui, B. M. W. and Barrett, H. H., Noise properties of the EM algorithm. II. Monte Carlo simulations *Phys Med Biol* **39:** 847-871 (1994).
21. Beekman F. J., Slijpen, E. T. and Niessen, W. J., Selection of task-dependent diffusion filters for the post–processing of SPECT images. *Phys Med Biol* **43:** 1713-1730 (1998).
22. Geman S. and McClure, D. E., Statistical methods for tomographic image reconstruction. *Bull Int Stat Inst* **52-4:** 55-21 (1987).
23. Mumcuoglu E. U., Leahy, R. M. and Cherry, S. R., Bayesian reconstruction of PET images: methodology and performance analysis. *Phys Med Biol* **41:** 1777-1807 (1996).

24. de Pierro A. R., On the convergence of an EM-type algorithm for penalized likelihood estimation in emission tomography. *IEEE Trans Med Imaging* **14:** 762-765 (1995).
25. Nuyts J., Beque, D., Dupont, P. *et al.*, A concave prior penalizing relative differences for maximum-a-posteriori reconstruction in emission tomography. *IEEE Trans Nucl Sci* **49:** 56-60 (2002).
26. Green P. J., Bayesian reconstructions from emission tomography data using a modified EM algorithm. *IEEE Trans Med Imaging* **9:** 84-93 (1990).
27. Ahn S. and Fessler, J. A., Globally convergent image reconstruction for emission tomography using relaxed ordered subsets algorithms. *IEEE Trans Med Imaging* **22:** 613-626 (2003).
28. Slijpen E. T. P. and Beekman, F. J., Comparison of post-filtering and filtering between iterations for SPECT reconstruction. *IEEE Trans Nucl Sci* **46:** 2233-2238 (1999).
29. Nuyts J., On estimating the variance of smoothed MLEM images *IEEE Trans Nucl Sci* **49:** 714-721 (2002).
30. Silverman B. W., Jones, M. C., Wilson, J. D. *et al.*, A smoothed EM approach to indirect estimation problems, with particular reference to stereology and emission tomography. *J R Statist Soc B* **52:** 271-324 (1990).
31. Mustafovic S., Thielemans, K., Hogg, D. *et al.*, ''Object dependency of resolution and convergence rate in OSEM with filtering'' Nuclear Science Symposium Conference Record, 2001 IEEE, Vol. 3; *pp* 1786-1790 (2001).
32. Fessler J. A. and Rogers, W. L., Spatial resolution properties of penalized-likelihood image reconstruction: space-invariant tomographs. *IEEE Trans Imag Proc* **5:** 1346-1358 (1996).
33. Nuyts J. and Fessler, J. A., A penalized-likelihood image reconstruction method for emission tomography, compared to postsmoothed maximum-likelihood with matched spatial resolution. *IEEE Trans Med Imaging* **22:** 1042-1052 (2003).
34. Stayman J. W. and Fessler, J. A., Compensation for nonuniform resolution using penalized-likelihood reconstruction in space-variant imaging systems. *IEEE Trans Med Imaging* **23:** 269-284 (2004).
35. Johnson V. E., Wong, W. H., Hu, X. *et al.*, Image restoration using Gibbs priors: boundary modeling, treatment of blurring, and selection of hyperparameter. *IEEE Trans Pattern Anal Machine Intell* **13:** 413-425 (1991).
36. Gindi G., Lee, M., Rangarajan, A. *et al.*, Bayesian reconstruction of functional images using anatomical information as priors. *IEEE Trans Med Imaging* **12:** 670-680 (1993).
37. Ardekani B. A., Braun, M., Hutton, B. F. *et al.*, Minimum cross-entropy reconstruction of PET images using prior anatomical information. *Phys Med Biol* **41:** 2497-2517 (1996).
38. Som S., Hutton, B. F. and Braun, M., Properties of minimum cross-entropy reconstruction of emission tomography with anatomically based prior. *IEEE Trans Nucl Sci* **45:** 3014-3021 (1998).
39. Rangarajan A., Hsiao, I.–T. and Gindi, G., A bayesian joint mixture framework for the integration of anatomical information in functional image reconstruction. *J Mat Imag Vision* **12:** 119-217 (2000).
40. Comtat C., Kinahan, P. E., Fessler, J. A. *et al.*, Clinically feasible reconstruction of 3D whole-body PET/CT data using blurred anatomical labels. *Phys Med Biol* **47:** 11-20 (2002).

41. Zaidi H., Montandon, M.-L. and Slosman, D. O., Magnetic resonance imaging-guided attenuation correction in 3D brain positron emission tomography. *Med Phys* **30:** 937-948 (2003).
42. Sastry S. and Carson, R. E., Multimodality Bayesian algorithm for image reconstruction in positron emission tomography: a tissue composition model. *IEEE Trans Med Imaging* **16:** 750-761 (1997).
43. Baete K., Nuyts, J., Van Paesschen, W. *et al*., "Anatomical based FDG-PET reconstruction for the detection of hypometabolic regions in epilepsy" *Nuclear Science Symposium Conference Record*, 2002 IEEE, Vol. 3; *pp* 1481-1485 (2002).
44. Defrise M., Kinahan, P. E., Townsend, D. W. *et al*., Exact and approximate rebinning algorithms for 3-D PET data. *IEEE Trans Med Imaging* **16:** 145-158 (1997).
45. Alenius S. and Ruotsalainen, U., Bayesian image reconstruction for emission tomography based on median root prior. *Eur J Nucl Med* **24:** 258-265 (1997).
46. Alenius S. and Ruotsalainen, U., Generalization of median root prior reconstruction. *IEEE Trans Med Imaging* **21:** 1413-1420 (2002).
47. Hsiao I. T., Rangarajan, A. and Gindi, G., A new convex edge-preserving median prior with applications to tomography. *IEEE Trans Med Imaging* **22:** 580-585 (2003).
48. Hudson H. M. and Larkin, R. S., Accelerated image reconstruction using ordered subsets of projection data. *IEEE Trans Med Imaging* **13:** 601-609 (1994).
49. Hutton B. F., Hudson, H. M. and Beekman, F. J., A clinical perspective of accelerated statistical reconstruction. *Eur J Nucl Med* **24:** 797-808 (1997).
50. Byrne C. L., Accelerating the EMML algorithm and related iterative algorithms by rescaled block-iterative methods. *IEEE Trans Imag Proc* **7:** 100-109 (1998).
51. Browne J. and de Pierro, A. B., A row-action alternative to the EM algorithm for maximizing likelihood in emission tomography. *IEEE Trans Med Imaging* **15:** 687-699 (1996).
52. Tanaka E. and Kudo, H., Subset-dependent relaxation in block-iterative algorithms for image reconstruction in emission tomography. *Phys Med Biol* **48:** 1405-1422 (2003).
53. Schmidlin P., Bellemann, M. E. and Brix, G., Subsets and overrelaxation in iterative image reconstruction. *Phys Med Biol* **44:** 1385-1396 (1999).
54. Lalush D. S., Frey, E. C. and Tsui, B. M., Fast maximum entropy approximation in SPECT using the RBI-MAP algorithm. *IEEE Trans Med Imaging* **19:** 286-294 (2000).
55. de Pierro A. R. and Yamagishi, M. E. B., Fast EM-like methods for maximum "a posteriori" estimates in emission tomography. *IEEE Trans Med Imaging* **20:** 280-288 (2001).
56. Leahy R. and Byrne, C. L., Recent developments in iterative image reconstruction for PET and SPECT. *IEEE Trans Med Imaging* **19:** 257-260 (2000).
57. Lange K. and Carson, R., EM reconstruction algorithms for emission and transmission tomography. *J Comput Assist Tomogr* **8:** 306-316 (1984).
58. Manglos S. H., Gagne, G. M., Krol, A. *et al*., Transmission maximum-likelihood reconstruction with ordered subsets for cone beam CT. *Phys Med Biol* **40:** 1225-1241 (1995).
59. Mumcuoglu E. U., Leahy, R., Cherry, S. R. *et al*., Fast gradient-based methods for Bayesian reconstruction of transmission and emission PET images. *IEEE Trans Med Imaging* **13:** 687-701 (1994).

60. Fessler J. A., Ficaro, E. P., Clinthorne, N. H. *et al.*, Grouped-coordinate ascent algorithms for penalized-likelihood transmission image reconstruction. *IEEE Trans Med Imaging* **16:** 166-175 (1997).
61. Nuyts J., Dupont, P., Stroobants, S. *et al.*, Simultaneous maximum a posteriori reconstruction of attenuation and activity distributions from emission sinograms. *IEEE Trans Med Imaging* **18:** 393-403 (1999).
62. Lalush D. S. and Tsui, B. M., Performance of ordered-subset reconstruction algorithms under conditions of extreme attenuation and truncation in myocardial SPECT. *J Nucl Med* **41:** 737-744 (2000).
63. Meikle S. R., Dahlbom, M. and Cherry, S. R., Attenuation correction using count-limited transmission data in positron emission tomography. *J Nucl Med* **34:** 143-150 (1993).
64. Tai Y.-C., Lin, K.-P., Dahlbom, M. *et al.*, A hybrid attenuation correction technique to compensate for lung density in 3-D total body PET. *IEEE Trans Nucl Sci* **43:** 323-330 (1996).
65. Xu M., Cutler, P. and Luk, W., An adaptive local threshold segmented attenuation correction method for whole-body PET imaging. *IEEE Trans Nucl Sci* **43:** 331-336 (1996).
66. Zaidi H., Diaz-Gomez, M., Boudraa, A. E. *et al.*, Fuzzy clustering-based segmented attenuation correction in whole-body PET imaging. *Phys Med Biol* **47:** 1143-1160 (2002).
67. Yavuz M. and Fessler, J. A., Statistical image reconstruction methods for randoms-precorrected PET scans. *Med Image Anal* **2:** 369-378 (1998).
68. Nuyts J., Michel, C. and Dupont, P., Maximum-likelihood expectation-maximization reconstruction of sinograms with arbitrary noise distribution using NEC-transformations. *IEEE Trans Med Imaging* **20:** 365-375 (2001).
69. Tsui B. M., Frey, E. C., Zhao, X. *et al.*, The importance and implementation of accurate 3D compensation methods for quantitative SPECT. *Phys Med Biol* **39:** 509-530 (1994).
70. Hutton B. F. and Lau, Y. H., Application of distance-dependent resolution compensation and post-reconstruction filtering for myocardial SPECT. *Phys Med Biol* **43:** 1679-1693 (1998).
71. Beekman F. J., Kamphuis, C. and Viergever, M. A., Improved SPECT quantitation using fully three-dimensional iterative spatially variant scatter response compensation. *IEEE Trans Med Imaging* **15:** 491-499 (1996).
72. Frey E. C. and Tsui, B. M. W., A practical method for incorporating scatter in a projector-backprojector for accurate scatter compensation in SPECT. *IEEE Trans Nucl Sci* **40:** 1107-1116 (1993).
73. Beekman F. J., de Jong, H. W. and van Geloven, S., Efficient fully 3-D iterative SPECT reconstruction with Monte Carlo-based scatter compensation. *IEEE Trans Med Imaging* **21:** 867-877 (2002).
74. Bowsher J. E. and Floyd, C. E., Jr., Treatment of Compton scattering in maximum-likelihood, expectation-maximization reconstructions of SPECT images. *J Nucl Med* **32:** 1285-1291 (1991).
75. de Jong H. W., Beekman, F. J., Viergever, M. A. *et al.*, Simultaneous (99m)Tc/(201)Tl dual-isotope SPET with Monte Carlo-based down-scatter correction. *Eur J Nucl Med Mol Imaging* **29:** 1063-1071 (2002).

76. Hudson H. M., Hutton, B. F., Larkin, R. S. *et al*., Investigation of multiple energy reconstruction in SPECT using MLEM. *J Nucl Med* **37:** 171P (abstract) (1996).
77. Carson R. E. and Lange, K., The EM parametric image reconstruction algorithm. *J Am Statist Assoc* **80:** 20-22 (1985).
78. Lalush D. S. and Tsui, B. M., Block-iterative techniques for fast 4D reconstruction using a priori motion models in gated cardiac SPECT. *Phys Med Biol* **43:** 875-886 (1998).
79. Matej S. and Browne, J. A., ''Performance of a fast maximum likelihood algorithm for fully 3D PET reconstruction'' in: *Three-Dimensional Image Reconstruction in Radiology and Nuclear Medicine*, edited by P. Grangeat and J-L. Amans Kluwer Academic Publishers, Dordrecht, (1996) *pp* 297-315.
80. Liu X., Comtat, C., Michel, C. *et al*., Comparison of 3-D reconstruction with 3D-OSEM and with FORE+OSEM for PET. *IEEE Trans Med Imaging* **20:** 804-814 (2001).
81. Fulton R. R., Hutton, B. F., Braun, M. *et al*., Use of 3D reconstruction to correct for patient motion in SPECT. *Phys Med Biol* **39:** 563-574 (1994).
82. Hutton B. F., Kyme, A. Z., Lau, Y. H. *et al*., A hybrid 3-D reconstruction/registration algorithm for correction of head motion in emission tomography. *IEEE Trans Nucl Sci* **49:** 188-194 (2002).
83. Bloomfield P. M., Spinks, T. J., Reed, J. *et al*., The design and implementation of a motion correction scheme for neurological PET. *Phys Med Biol* **48:** 959-978 (2003).
84. Barrett H. H., White, T. and Parra, L. C., List-mode likelihood. *J Opt Soc Am A* **14:** 2914-2923 (1997).
85. Reader A. J., Erlandsson, K., Flower, M. A. *et al*., Fast accurate iterative reconstruction for low-statistics positron volume imaging. *Phys Med Biol* **43:** 835-846 (1998).
86. Matej S. and Lewitt, R. M., Practical considerations for 3-D image reconstruction using spherically symmetric volume elements. *IEEE Trans Med Imaging* **15:** 68-78 (1996).
87. Swensson R. G., Unified measurement of observer performance in detecting and localizing target objects on images. *Med Phys* **23:** 1709-1725 (1996).
88. Boellaard R., van_Lingen, A. and Lammertsma, A. A., Experimental and clinical evaluation of iterative reconstruction (OSEM) in dynamic PET: quantitative characteristics and effects on kinetic modeling. *J Nucl Med* **42:** 808-817 (2001).
89. Montandon M.-L., Slosman, D. O. and Zaidi, H., Assessment of the impact of model-based scatter correction on ^{18}F-[FDG] 3D brain PET in healthy subjects using statistical parametric mapping. *Neuroimage* **20:** 1848-1856 (2003).
90. Mesina C. T., Boellaard, R., Jongbloed, G. *et al*., Experimental evaluation of iterative reconstruction versus filtered backprojection for 3D [15O]water PET activation studies using statistical parametric mapping analysis. *Neuroimage* **1:** 1170-1179 (2003).
91. Beekman F. J. and Viergever, M. A., Fast SPECT simulation including object shape dependent scatter. *IEEE Trans Med Imag* **14:** 271-282 (1995).
92. Ljungberg M., Strand, S.-E. and King, M. A., ''Monte Carlo calculations in nuclear medicine: Applications in diagnostic imaging'' (Institute of Physics Publishing, Bristol, 1998).

93. Zaidi H. and Sgouros, G., "Therapeutic applications of Monte Carlo calculations in nuclear medicine" (Institute of Physics Publishing, Bristol, 2002).
94. Barrett H. H., Denny, J. L., Wagner, R. F. *et al.*, Objective assessment of image quality. II. Fisher information, Fourier crosstalk, and figures of merit for task performance. *J Opt Soc Am A* **12:** 834-852 (1995).
95. Metz C. E., ROC methodology in radiologic imaging. *Invest Radiol* **21:** 720-733 (1986).
96. Barrett H. H., Abbey, C. K. and Clarkson, E., Objective assessment of image quality. III. ROC metrics, ideal observers, and likelihood-generating functions. *J Opt Soc Am A Opt Image Sci Vis* **15:** 1520-1535 (1998).
97. Gifford H. C., King, M. A., de_Vries, D. J. *et al.*, Channelized hotelling and human observer correlation for lesion detection in hepatic SPECT imaging. *J Nucl Med* **41:** 514-521 (2000).
98. Lartizien C., Kinahan, P. E., Swensson, R. *et al.*, Evaluating image reconstruction methods for tumor detection in 3-dimensional whole-body PET oncology imaging. *J Nucl Med* **44:** 276-290 (2003).
99. Lonneux M., Borbath, I., Bol, A. *et al.*, Attenuation correction in whole-body FDG oncological studies: the role of statistical reconstruction. *Eur J Nucl Med* **26:** 591591-598 (1999).
100. Blocklet D., Seret, A., Popa, N. *et al.*, Maximum-likelihood reconstruction with ordered subsets in bone SPECT. *J Nucl Med* **40:** 19781978-1984 (1999).
101. Hatton R., Hutton, B. F. and Angelides, S., Improved tolerance to missing data in myocardial perfusion SPECT using OSEM reconstruction. *Eur J Nucl Med Mol Imaging* **31:** 857-861 (2004).
102. Riddell C., Carson, R. E., Carrasquillo, J. A. *et al.*, Noise reduction in oncology FDG PET images by iterative reconstruction: a quantitative assessment. *J Nucl Med* **42:** 1316-1323 (2001).
103. Fulton R. R., Meikle, S. R., Eberl, S. *et al.*, Signal to noise ratio in neurological FDG PET: Comparison of FBP and OS-EM reconstruction algorithms. *J Nucl Med* **37:** 220P (abstract) (1996).
104. Visvikis D., Cheze_LeRest, C., Costa, D. C. *et al.*, Influence of OSEM and segmented attenuation correction in the calculation of standardised uptake values for [18F]FDG PET. *Eur J Nucl Med* **28:** 1326-1335 (2001).
105. Iida H. and Eberl, S., Quantitative assessment of regional myocardial blood flow with thallium-201 and SPECT. *J Nucl Cardiol* **5:** 313-331 (1998).
106. Friston K., Holmes, A., Worsley, K. *et al.*, Statistical parametric maps in functional imaging: a general linear approach. *Human Brain Mapping* **2:** 189-210 (1995).
107. Braem A., Chamizo Llatas, M., Chesi, E. *et al.*, Feasibility of a novel design of high-resolution parallax-free Compton enhanced PET scanner dedicated to brain research. *Phys Med Biol* **49:** 2547-2562 (2004).
108. Buvat I., Lazaro, D. and Breton, V., "Fully 3D Monte Carlo reconstruction in SPECT: proof of concept and is that worthwhile?" *Conf. Proc. of the VIIth International Meeting on Fully Three-dimensional Image Reconstruction in Radiology and Nuclear Medicine*, 29 June-4 July 2003, Saint-Malo, France. Available on CD-ROM.